UNIVERSITY OF HULL

ENHANCING PNEUMONIA DIAGNOSIS THROUGH CUSTOM CNN MODELS AND IMBALANCE TREATMENT

Uzochukwu Ogbonna[a], Eva Sousa[b]

Department of Science and Engineering, University of Hull, Hull, United Kingdom

Abstract

Pneumonia, a global health concern impacting both the young and adults, necessitates precise diagnostics due to its substantial morbidity and mortality rates. Timely pneumonia detection using X-ray analysis and deep learning methods plays a vital role in disease management. However, challenges persist regarding interpretability and the quality of datasets. This study addresses the critical challenge of class imbalance in medical image datasets for pneumonia detection via X-ray images. Employing a custom convolutional neural network (CNN) model, we integrate innovative under-sampling and oversampling methods on a dataset of 5,232 X-ray images stratified into two classes (Class 0 = 3,883 and Class 1 = 1,349). The under-sampling of Class 0 in the initial phase rectifies class imbalance, while the subsequent oversampling of Class 1 ensures balance. The Under-sampled class exhibits 82% accuracy with discernible signs of overfitting. However, the best accuracy is obtained on the Oversampled state with a value of 88%. Hyperparameter tuning was also implemented on convolution blocks, learning rate, number of epochs, and optimizers to determine the best and most accurate method. This holistic approach highlights the importance of addressing class imbalance intricacies and optimizing hyperparameters for efficacious pneumonia detection through the adept deployment of CNN models.

Keywords: **Pneumonia, X-ray Analysis, Deep Learning, Class Imbalance, Under-sampling, Oversampling, CNN Models, Disease Management, Diagnostic Precision, Chest Xray images**

1.0 Introduction

Pneumonia is a serious lung infection that affects both kids and adults worldwide. It is a major cause of illness and death, especially in children under five years old(1) (2) . The World Health Organization (WHO) says about 15% of deaths in these young kids are due to pneumonia, and it also poses risks for adults, especially older people, and those with weakened immune systems(3) (4) (5). Diagnosing pneumonia is crucial because it's widespread, has various causes, and imposes a significant global health burden, especially on vulnerable populations. Swift and accurate diagnosis is key for early treatment, reducing complications, and improving patient outcomes(6) (7) (8). The importance of pneumonia diagnosis has been emphasized even more with its global impact and its connection to other health concerns, like

COVID-19 (9) . Getting the right diagnosis is not only important for successful treatment but also for reducing mortality rates and addressing public health issues (10) (11).

In medical imaging, X-rays play a crucial role in diagnosing pneumonia. X-ray imaging uses radiation to create detailed lung images, showing changes in tissue density that can reveal pneumonia-related abnormalities (12) (13). X-rays are fundamental in pneumonia diagnosis because they can detect subtle structural changes in the lungs (14) (15) (16). Convolutional Neural Networks (CNNs), a type of deep learning, have become essential tools in classifying medical images, especially X-rays (17) (18) (19) (20). CNNs are great at automatically learning and extracting complex features from images, making them ideal for recognizing patterns in medical images (21) (22) (23) . In pneumonia diagnosis, CNNs have shown remarkable abilities in distinguishing between healthy and affected lung tissues, improving diagnostic accuracy (24) (25).

Research in pneumonia detection from chest X-ray images has drawn considerable attention, with a focus on cutting-edge models and diagnostic methods rooted in deep-learning architecture (26) (27). One notable contribution is exemplified by Rajpurkar and colleagues, who introduced a CNN algorithm featuring 121 layers tailored for pneumonia detection(28). Kermany et al. proposed a deep learning model for medical disease diagnosis, achieving a commendable accuracy of 92.8%, comparable to human specialists(29). However, this model, while enhancing performance, relies on lower-level features. Saraiva et al. focused on childhood pneumonia classification using a CNN with 7 convolutional layers and 3 dense layers, Nevertheless, the study acknowledged limitations related to a small number of convolutional layers and restricted data size (30). In another approach, Wu et al. introduced a hybrid model combining an adaptive median filter CNN and random forest for pneumonia prediction, achieving high accuracy but requiring additional preprocessing due to the adaptive median filter(31).

Despite these advancements, challenges persist, particularly in addressing class imbalance within medical image datasets for pneumonia detection. This study builds upon these works, specifically focusing on the crucial aspect of class imbalance in pneumonia detection from X-ray images. The novelty here lies in a thorough investigation into two methods for treating class imbalance: under-sampling and oversampling. The study aims to compare the effectiveness of these techniques in handling imbalance for pneumonia detection in chest X-ray images using Convolutional Neural Networks (CNNs). By doing so, this research seeks to provide valuable insights, contribute to addressing existing challenges, and ultimately enhance the accuracy and efficiency of pneumonia diagnostics. Subsequent sections detail the methodology, experimental results, discussions and the overall contribution to evolving pneumonia diagnostics. Below is the architectural framework of the CNN model

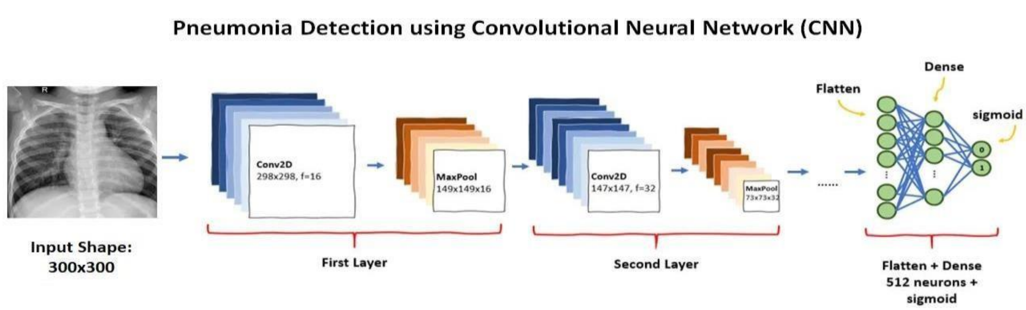

Fig 1 . CNN Architecture (32)

Convolutional operations within CNNs architecture **Fig 1** above, often referred to as **convolution blocks**, are employed for feature extraction. **Max pooling** serves as a down-sampling technique, reducing spatial dimensions while preserving crucial information. **The Flatten** operation transforms data into a 1-D array, **while Dense layers** establish connections between neurons for classification purposes. Finally, the Sigmoid activation function yields binary output probabilities, ranging from 0 to 1.

2.0 Methodology

2.1 Data Preprocessing

The dataset for this study was obtained from a Kaggle open-source repository(33), consisting of chest X-ray images categorized into two classes: Healthy X-ray images and Pneumonia X-ray images. The dataset was already partitioned into training and testing subset. An initial examination of the dataset highlighted an inherent imbalance between the two where the Class 0 had a total of 3883 X-ray image samples of Pneumonia Images while Class 1 have a total of 1349 of Healthy image samples classes.

To rectify the class imbalance, two distinct approaches were implemented. In the first instance, imbalanced random sampling techniques from the imblearn library were employed, specifically using under-sampling to address the over-represented Class 0 and bring it to balance with Class 1. This step resulted in the two classes having a total of 1349 X-ray images each. Conversely, in the second instance, oversampling techniques were also utilized to address the under-represented Class 1, achieving a balanced distribution with Class 0 with each class having a total of 3883. Furthermore, series of preprocessing steps were applied to the images.

- **Image Resizing**

The dimensionality of the images was reduced through resizing to achieve a uniform input size of 256x256 pixels. This resizing was undertaken to enhance model compatibility, computational efficiency and retain the original features of the X-ray images.

- **Normalization**

Normalization process involves dividing each pixel value by 256 to rescale them to a range between 0 and 1. This ensures that the pixel values are within a standardized range suitable for machine learning models. The original values of these images are divided by 256 to obtain the corresponding normalized pixel values. This normalization process is crucial for our models as it facilitates better convergence during training and ensures that the model is less sensitive to the magnitude of pixel values in the input images(34).

- **Augmentation**

Augmentation techniques, such as rotation (30 degrees), horizontal flipping, brightness adjustment, and zooming (20%), were exclusively applied to the training dataset to artificially expand its size. Additionally, random horizontal and vertical shifts (10% of image width and height, respectively) were introduced, providing spatial variations. Shearing transformations (0.0) and ZCA whitening (False with epsilon set to 1e-6) were not applied. Channel shifts (0.0) for color variations, constant pixel value filling (cval=0.0), and "nearest" fill mode was used during augmentation. Below are some of the samples of the augmented images used to enhance our training dataset

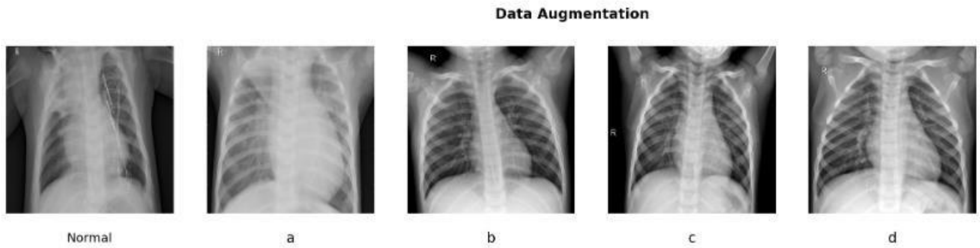

Fig 2. Data Augmentation: **(a)** brightness **(b)** width increase **(c)** random rotation **(d)** zoom

2.2 Model Design

A custom Convolutional Neural Network (CNN) architecture was meticulously designed for pneumonia detection in X-ray images using the Sequential model. The architecture started with an initial Conv2D layer featuring 16 filters and a 3x3 kernel, applying Rectified Linear Unit (ReLU) activation to process input images of size (256, 256, 1). Subsequently, a MaxPooling2D layer with a 2x2 pool size enhances feature extraction while reducing spatial dimensions.

Continuing the network progression, another Conv2D layer introduces 32 filters with ReLU activation, followed by a MaxPooling2D layer. A Dropout layer is strategically added with a dropout rate of 0.5, contributing to regularization by randomly deactivating neurons during training. The architecture then incorporates another Conv2D layer, this time with 64 filters and ReLU activation, followed by a MaxPooling2D layer. A Dropout layer (0.6) is introduced for further regularization.

Additional Convolutional Blocks amplify the feature extraction process, including a Conv2D layer with 128 filters and ReLU activation, followed by a MaxPooling2D layer. Subsequently, a Conv2D layer with 256 filters and ReLU activation, followed by another MaxPooling2D layer, is introduced. A Dropout layer (0.5) further contributes to regularization. The model then features another Conv2D layer with 128 filters and ReLU activation, followed by a MaxPooling2D layer. A Dropout layer (0.6) is applied, providing additional regularization.

The architecture transitions to a Flatten layer, transforming the 3D output from the convolutional layers into a 1D array, preparing the data for subsequent Dense layers. Two Dense layers follow: the first with 128 neurons and ReLU activation incorporates L2 regularization (kernel_regularizer=regularizes. l2(2e-0)) to prevent overfitting. A Dropout layer (0.6) follows, contributing to the model's regularization. The final Dense layer, with a single neuron and a sigmoid activation function, facilitates binary classification, making it apt for the pneumonia detection task.

The incorporation of dropout, L2 regularization, and a systematic hierarchy of convolutional and pooling layers ensures a well-balanced architecture. This thoughtful combination optimizes feature extraction, introduces regularization, and enables effective classification for accurate pneumonia detection from X-ray images. Furthermore, the learning rate for the optimizer was set to 0.00001. The Adam optimizer was initialized with this specified learning rate. Finally, the model was compiled using the Adam optimizer, binary cross entropy loss function, and accuracy as the evaluation metric.

2.3 Training and Validation

During the training phase, the two Convolutional Neural Network (CNN) underwent a process of optimization using the designated training dataset. This process involved iterative adjustments to the network's internal parameters to minimize a predefined loss function. Simultaneously, the model's performance on a separate validation dataset (Test dataset) was meticulously monitored. This vigilant observation of the validation set was crucial to detect signs of overfitting, where the model might excel in learning the training data but struggle to generalize to unseen data. By scrutinizing the validation performance, I ensured that the CNN achieved robust and generalized results, aligning with the objective of effective pneumonia detection in X-ray images.

2.4 Evaluation and Interpretation

During the Evaluation and Interpretation phase, evaluated the two CNN model’s effectiveness in pneumonia detection. Employed established evaluation metrics like Training and validation and Training and validation accuracy, accuracy, precision, recall and F1-score and confusion matrix to comprehensively gauge the separate performance of each model.

2.5 Model Comparison

This phase involves an extensive assessment of two custom-designed CNN models for pneumonia detection in X-ray images. The aims are to determine which model performs better across key metrics which include Training and validation, Training and validation accuracy, Model accuracy, precision, recall and F1-score and confusion matrix. By scrutinizing these metrics, we discern the superior performer in distinguishing between Healthy and Pneumonia X-ray images.

3.0 Results

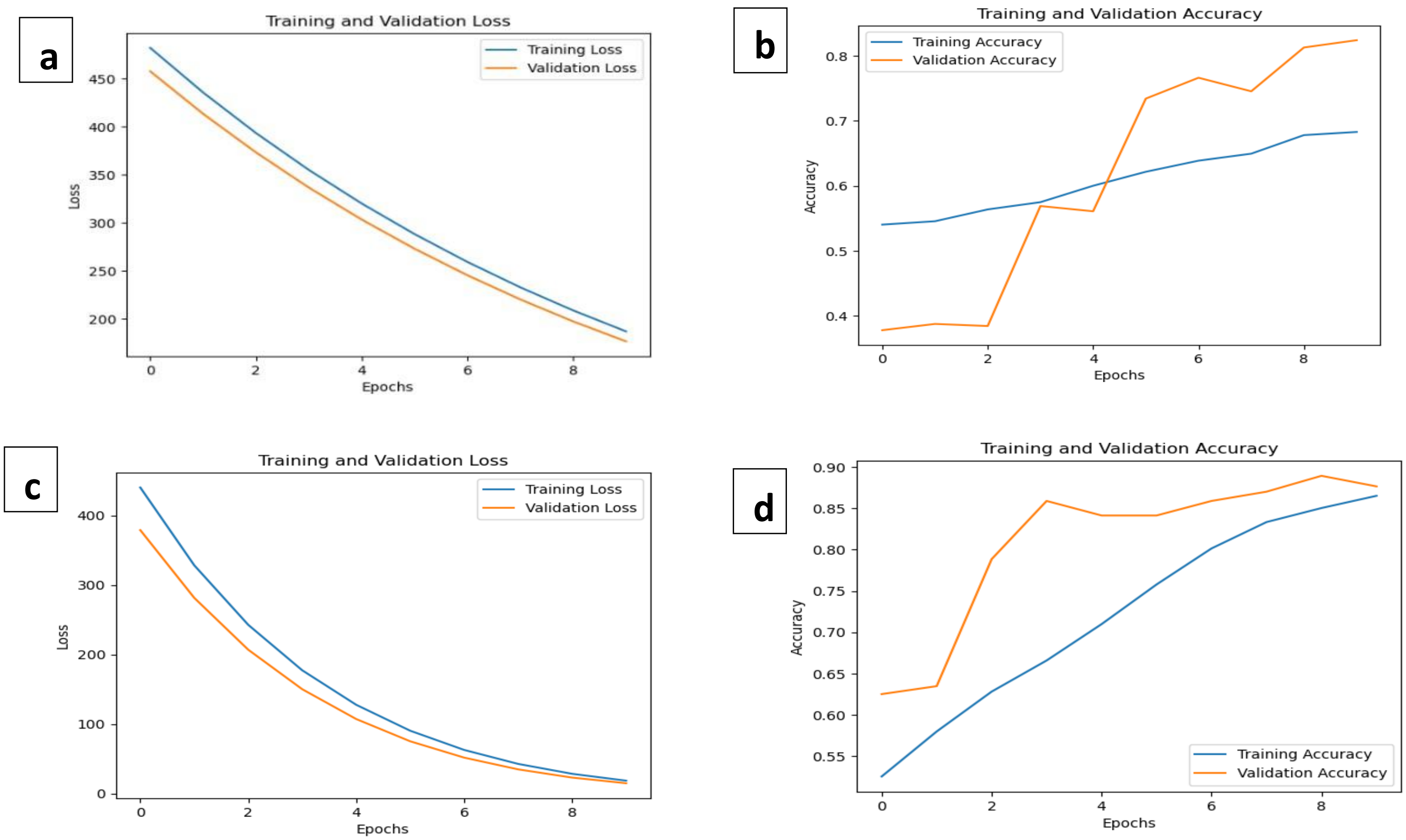

Fig 3. (a) Undersampling- Training and validation loss **(b)** Undersampling- Training and validation accuracy **(c)** Oversampling- Training and validation loss **(d)** Oversampling- Training and validation accuracy

In a thorough examination of the training dynamics spanning 10 epochs, insightful trends emerged from the comparative analysis of both models. The Under-sampling model started at 54.04% training accuracy and ended at 68.27%, as depicted in **Fig 3. (a)**. However, the validation accuracy exhibited only a marginal increase, suggesting potential overfitting and culminating in a final value of 82.37%, as illustrated in **Fig 3. (b)**. In contrast, the Oversampling model showed more improvement, starting at 52.52% and reaching an impressive 86.52%, as showed in **Fig 3. (c)**. Notably, the validation accuracy consistently increased, reflecting enhanced generalization and stability, ultimately peaking at an impressive 87.66%, as depicted in **Fig 3. (d)**. Turning our attention to the loss metrics, the Under-sampling model experienced a modest reduction in training loss from 439.98 to 18.42 and validation loss from 378.71 to 14.76, as can be seen in **Fig 3. (a)**. Conversely, the Oversampling model demonstrated a substantial decrease in training loss from 482.27 to 186.99 and in validation loss from 457.83 to 176.59, evident in **Fig 3. (c)**. The consistent decline in loss metrics for the Oversampling model underscores superior convergence and enhanced generalization compared to the Under-sampling model.

Table 1: Under sampling – Classification report

	Precision	recall	f1-score	Support
Pneumonia (Class 0)	0.87	0.84	0.86	390
Normal (Class 1)	0.75	0.79	0.77	234
Accuracy			**0.82**	624
macro avg	0.81	0.82	0.81	624
weighted avg	0.83	0.82	0.82	624

Table 2: Oversampling – Classification report

	Precision	recall	f1-score	Support
Pneumonia (Class 0)	0.86	0.96	0.91	390
Normal (Class 1)	0.92	0.73	0.82	234
Accuracy			**0.88**	624
macro avg	0.89	0.85	0.86	624
weighted avg	0.88	0.88	0.87	624

The oversampling classification report in **Table 2** further illuminates a well-balanced performance, with precision and recall scores exhibiting equilibrium for both classes. For Pneumonia (Class 0), the precision is 0.86, recall is 0.96, and the F1-score is 0.91. Similarly, for Healthy (Class 1), the precision is 0.92, recall is 0.73, and the F1-score is 0.82. The overall accuracy stands at an impressive 88%, with both macro and weighted averages consistently at 86%, highlighting the robustness of the model. Contrastingly, the under-sampling classification report in **Table 1** exposes a less balanced performance, marked by imbalances in precision and recall scores for both classes. For Pneumonia (Class 0), the precision is 0.87, recall is 0.84, and the F1-score is 0.86. For Normal (Class 1), the precision is 0.75, recall is 0.79, and the F1-score is 0.77. The overall accuracy is 82%, with macro and weighted averages at 81%, underscoring a less balanced performance compared to the oversampling approach.

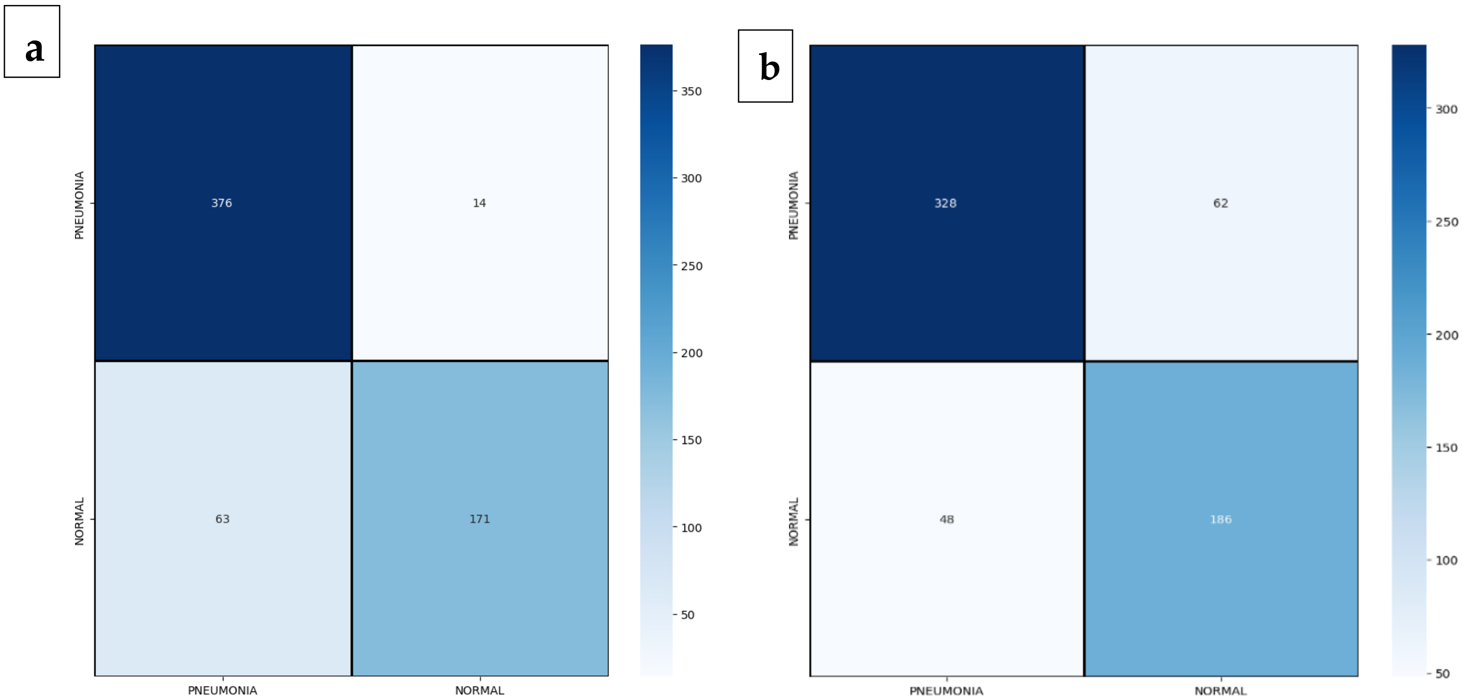

Fig 4. **Confusion Matrix:** (**a) Oversampling** (**b) Under sampling**

Examining the confusion matrices, the oversampling model in **Fig 4. (a)** predicted 376 True Negatives (Pneumonia) and 14 False Positives (Healthy) for Class 0, while predicting 63 False Negatives (Pneumonia) and 171 True Positives (Healthy) for Class 1. This demonstrates superior performance compared to the undersampling model in **Fig 4. (b)**, which showed 328 True Negatives (Pneumonia), 62 False Positives (Healthy) for Class 0, and 48 False Negatives (Pneumonia), 186 True Positives (Healthy) for Class 1. For the undersampling model, it reveals a sensitivity of 84% for detecting pneumonia (Class 0) and specificity of 75% for identifying normal cases (Class 1), indicating a potential class imbalance as noted in Table 1. On the other hand, the oversampling model showcases a well-balanced confusion matrix, maintaining high sensitivity (87.5%) and specificity (97.03%) for both pneumonia and normal classes (Class 0 and Class 1) as seen in Table 1. It boasts an improved precision of 0.78 and recall of 0.88 for pneumonia, manifesting effective classification and reduced bias.

In summary, in the course of the 10 epochs, the oversampling model demonstrates superior stability, accuracy improvement, and reduced loss compared to the undersampling model. The all-inclusive assessment, incorporating the confusion matrix, accuracy, precision, recall, and training and validation loss, collectively underlines the effectiveness and robustness of the oversampling approach in handling imbalanced data for pneumonia detection in X-ray images.

4.0 Discussion

In this research, Sampling techniques and CNN was employed on unbalanced X-ray dataset. Recognizing the importance of balancing the dataset for CNN applications is crucial for robust model performance. This is the reason why data imbalances were addressed using both undersampling and oversampling methodologies, aiming to enhance the model's accuracy and mitigate the risk of overfitting—common challenges in unbalanced datasets. The utilization of CNN on balanced data ensures a more robust and accurate classification, contributing to the effectiveness of the model outcome.

In evaluating the CNN models, a comprehensive comparison of under-sampling and oversampling techniques was conducted, unveiling distinct trends in model effectiveness. Similar to the study by Buda et al. on CNNs for digit recognition (35), our findings revealed that oversampling emerged as a critical and effective strategy, demonstrating unparalleled performance. The oversampling technique showcased exceptional precision, recall, and F1-score for both Pneumonia (Class 0) and Normal (Class 1), mirroring the trends observed in the digit recognition study. Specifically, our model achieved a precision of 0.86, recall of 0.96, and F1-score of 0.91 for Pneumonia, emphasizing how oversampling effectively balances precision and recall as seen in **Table 2**. This aligns with the notion that oversampling generally performs well under certain conditions, as observed in the digit recognition study (35), where it excelled as the number of minority classes and the imbalance ratio increased. Oversampling has also been reported to be in reference to classification performance (36).

Moreover, the substantial decrease in training loss from 482.27 to 186.99 and validation loss from 457.83 to 176.59, as evident in **Fig 3. (c)** and **Fig 3. (d)**, indicated the superior convergence and enhanced generalization capabilities of our model, similar to the observations made in the digit recognition study. The consistent decline in these loss metrics emphasized the model's robustness and its ability to make accurate predictions on unseen data, aligning with the importance of enhanced generalization in both studies.

In our study, the impressive overall accuracy of 88% in our model can be attributed to the detailed duplication of minority class instances during the oversampling process. By strategically augmenting the dataset, the oversampling technique ensured greater exposure to distinctive features within the minority class, mitigating bias towards the majority class. This approach significantly contributed to the superior performance of the model, reflecting its capacity to make well-informed and balanced predictions across both classes. This finding aligns with the broader literature on addressing class imbalance in machine learning, as discussed in previous studies (37) (38) (39). The effectiveness of oversampling in mitigating class imbalance and improving model performance has been widely recognized (40) (41). It is important to note that while oversampling proved effective in our specific model and dataset, the choice of imbalance mitigation method may depend on the characteristics of the dataset and the specific machine learning algorithm employed.

The under-sampling model, starting with an initial training accuracy of 54.04% and concluding at 68.27%, showed a notable improvement over the epochs, as depicted in **Fig 3. (b)**. However, the validation accuracy exhibited only a marginal increase, reaching 82.37%, hinting at potential overfitting in **Fig 3. (b)** and **Table 1** despite augmenting the dataset**.** The modest reduction in training loss from 439.98 to 18.42 and validation loss from 378.71 to 14.76 in **Fig 3. (a)** reflects the model's attempt to optimize, but the limitations of the under-sampling methodology became apparent. The under-sampling methodology, involving the systematic removal of instances from the majority class, posed challenges leading to imbalances in precision and recall, particularly evident in Class 1, seen in **Table 1**. This imbalance contributed to a lower overall accuracy of 82%, accompanied by manifestations of overfitting, shedding light on the limitations of under-sampling in achieving a balanced model for pneumonia detection. The reduction in the dataset size might have led to the loss of crucial information, impacting the model's ability to generalize effectively.

Our study diverged from the findings of the Physical activity classification when it came to the under-sampling approach. Unlike the Physical activity classification study in (42), where under-sampling showed notable improvement better than oversampling though the model used was Random Forest classifier, while our pneumonia detection model experienced challenges with under sampling technique. This could mean that different machine learning model and dataset condition could influence the performance of this sampling techniques. In the Physical activity classification study, the dataset heterogeneity positively influenced performance, while in our study, the limited dataset size hinted at potential overfitting reflecting limitations in achieving balance and model robustness. The differences between our study and previous studies emphasizes the importance of dataset characteristics in model training.

However, our findings align with previous studies that emphasize the challenges associated with under-sampling in achieving balance and model robustness (43), (44),(45). While the under-sampling model demonstrated improvement, the oversampling approach outperformed in terms of both accuracy and stability, showcasing a more effective strategy for handling imbalanced data in pneumonia detection models.

Our study-constrained dataset size of 5232 remains a challenge, impeding the model from reaching its full potential. While the adoption of oversampling techniques successfully addressed class imbalance, the limited dataset size hinted at potential overfitting if the training duration was prolonged. Thus, a larger, more diverse dataset meticulously curated to maintain a balanced representation of classes would undoubtedly fortify the model's capacity to generalize and make robust predictions across diverse scenarios. The challenges encountered in selecting the optimal resampling technique underscore the complexities inherent in this critical decision.

Additionally, an exploration of alternative machine learning algorithms is important to improve the model's accuracy, interpretability, and efficiency in medical image analysis. The integration of multimodal features, incorporating clinical and demographic data, holds immense potential for crafting a more comprehensive model, poised to revolutionize pneumonia detection in medical imaging.

5.0 Conclusion:

In summary, spanning 10 epochs, the Oversampling model not only demonstrated superior stability, accuracy improvement, and reduced loss compared to the Under-sampling model but also aligns with the current discourse in the literature. The holistic assessment, incorporating confusion matrix, accuracy, precision, recall, and training and validation loss, collectively underscores the effectiveness and robustness of the Oversampling approach in handling imbalanced data for pneumonia detection in X-ray images.

These findings contribute to the evolving understanding of effective strategies for model development in the domain of medical image classification.

Reference

1. What is pneumonia in children? | Asthma + Lung UK [Internet]. 2022 [cited 2023 Dec 11]. Available from: https://www.asthmaandlung.org.uk/conditions/pneumonia/child/what-is

2. Acute Lower Respiratory Tract Infection|Diagnosis & Treatment|ELF [Internet]. European Lung Foundation. [cited 2023 Dec 11]. Available from: https://europeanlung.org/en/information-hub/lung-conditions/acute-lower-respiratory-infections/

3. Pneumonia [Internet]. [cited 2023 Aug 17]. Available from: https://www.who.int/health-topics/pneumonia

4. Dadonaite B, Roser M. Pneumonia. Our World in Data [Internet]. 2023 Dec 11 [cited 2023 Dec 11]; Available from: https://ourworldindata.org/pneumonia

5. ‘It can get very serious, very quickly’: the symptoms, risks, and ways to prevent pneumonia everyone should know about | GSK [Internet]. [cited 2023 Dec 11]. Available from: http://www.gsk.com/en-gb/behind-the-science-magazine/pneumonia-symptoms-risks-vaccines-treatment/

6. How is pneumonia diagnosed and treated? | Asthma + Lung UK [Internet]. 2022 [cited 2023 Dec 11]. Available from: https://www.asthmaandlung.org.uk/conditions/pneumonia/diagnosis

7. Cleveland Clinic [Internet]. [cited 2023 Dec 11]. Pneumonia: Causes, Symptoms, Diagnosis & Treatment. Available from: https://my.clevelandclinic.org/health/diseases/4471-pneumonia

8. Kallander K, Burgess DH, Qazi SA. Early identification and treatment of pneumonia: a call to action. Lancet Glob Health. 2015 Nov 12;4(1):e12–3.

9. Bernstein S. WebMD. [cited 2023 Dec 11]. Coronavirus and Pneumonia. Available from: https://www.webmd.com/covid/covid-and-pneumonia

10. Why is it important to diagnose pneumonia? - Ṣíṣe àwárí ni Google [Internet]. [cited 2023 Dec 11]. Available from: https://www.google.com/search?sca_esv=589872414&sxsrf=AM9HkKlP0vUTRjDFNpRbKw8FE7hmr5eYxg:1702326150220&q=Why+is+it+important+to+diagnose+pneumonia%3F&sa=X&ved=2ahUKEwiT6eaLm4iDAxXZXEEAHYXSCiAQzmd6BAgbEAY&biw=1707&bih=813&dpr=0.8

11. Healthline [Internet]. 2018 [cited 2023 Dec 11]. Can You Die from Pneumonia? What You Need to Know. Available from: https://www.healthline.com/health/pneumonia/can-you-die-from-pneumonia

12. Tsakok MT, Gleeson FV. The chest radiograph in heart disease. Medicine. 2018 Aug;46(8):453–7.

13. Radiology (ACR) RS of NA (RSNA) and AC of. Radiologyinfo.org. [cited 2023 Dec 11]. chestrad. Available from: https://www.radiologyinfo.org/en/info/chestrad

14. Radiology (ACR) RS of NA (RSNA) and AC of. Radiologyinfo.org. [cited 2023 Dec 11]. Pneumonia. Available from: https://www.radiologyinfo.org/en/info/pneumonia

15. Tests for Lung Disease - Tests for Lung Disease | NHLBI, NIH [Internet]. [cited 2023 Dec 11]. Available from: https://www.nhlbi.nih.gov/health/lung-tests

16. Cleveland Clinic [Internet]. [cited 2023 Dec 11]. Chest X-ray (CXR): What You Should Know & When You Might Need One. Available from: https://my.clevelandclinic.org/health/diagnostics/10228-chest-x-ray

17. Singh Y, Farrelly C, Hathaway QA, Choudhary A, Carlsson G, Erickson B, et al. The Role of Geometry in Convolutional Neural Networks for Medical Imaging. Mayo Clinic Proceedings: Digital Health. 2023 Dec 1;1(4):519–26.

18. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging. 2018 Aug;9(4):611–29.

19. Convolutional neural networks in medical image understanding: a survey - PMC [Internet]. [cited 2023 Dec 11]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7778711/

20. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intel. 2022 Mar 1;15(1):1–22.

21. Taye MM. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. Computation. 2023 Mar;11(3):52.

22. Saha S. Medium. 2022 [cited 2023 Dec 11]. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. Available from: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

23. Introduction to Convolution Neural Network [Internet]. GeeksforGeeks. 2017 [cited 2023 Dec 11]. Available from: https://www.geeksforgeeks.org/introduction-convolution-neural-network/

24. Nessipkhanov D, Davletova V, Kurmanbekkyzy N, Omarov B. Deep CNN for the Identification of Pneumonia Respiratory Disease in Chest X-Ray Imagery. IJACSA [Internet]. 2023 [cited 2023 Dec 11];14(10). Available from: http://thesai.org/Publications/ViewPaper?Volume=14&Issue=10&Code=IJACSA&SerialNo=69

25. Motta PC, Cortez PC, Silva BRS, Yang G, de Albuquerque VHC. Automatic COVID-19 and Common-Acquired Pneumonia Diagnosis Using Chest CT Scans. Bioengineering (Basel). 2023 Apr 26;10(5):529.

26. Singh S, Rawat S, Gupta M, Tripathi B, Alanzi F, Majumdar A, et al. Deep Attention Network for Pneumonia Detection Using Chest X-Ray Images. Computers, Materials and Continua. 2022 Oct 6;74:1673–91.

27. Mabrouk A, Díaz Redondo RP, Dahou A, Abd Elaziz M, Kayed M. Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks. Applied Sciences. 2022 Jan;12(13):6448.

28. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning [Internet]. arXiv; 2017 [cited 2023 Dec 11]. Available from: http://arxiv.org/abs/1711.05225

29. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell. 2018 Feb 22;172(5):1122-1131.e9.

30. Saraiva A, Ferreira N, Sousa L, Carvalho da Costa N, Sousa J, Santos D, et al. Classification of Images of Childhood Pneumonia using Convolutional Neural Networks. In 2019. p. 112–9.

31. Mittal A, Kumar D, Mittal M, Saba T, Abunadi I, Rehman A, et al. Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images. Sensors. 2020 Jan;20(4):1068.

32. Thetechwriters. Pre-trained Stacked Model to Detect Pneumonia [Internet]. Analytics Vidhya. 2021 [cited 2023 Dec 11]. Available from: https://www.analyticsvidhya.com/blog/2021/12/pneumonia-detection-with-pre-trained-stacked-model/

33. Chest X-ray Images [Internet]. [cited 2023 Dec 11]. Available from: https://www.kaggle.com/datasets/tolgadincer/labeled-chest-xray-images

34. Ponraj A. A Tip A Day — Python Tip #8: Why should we Normalize image pixel values or divide by 255? [Internet]. Analytics Vidhya. 2021 [cited 2023 Dec 11]. Available from: https://medium.com/analytics-vidhya/a-tip-a-day-python-tip-8-why-should-we-normalize-image-pixel-values-or-divide-by-255-4608ac5cd26a

35. Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw. 2018 Oct;106:249–59.

36. Kim M, Hwang KB. An empirical evaluation of sampling methods for the classification of imbalanced data. PLoS One. 2022 Jul 28;17(7):e0271260.

37. Qu W, Balki I, Mendez M, Valen J, Levman J, Tyrrell PN. Assessing and mitigating the effects of class imbalance in machine learning with application to X-ray imaging. Int J CARS. 2020 Dec 1;15(12):2041–8.

38. Barandela R, Sánchez JS, Garcıa V, Rangel E. Strategies for learning in class imbalance problems. Pattern Recognition. 2003 Mar 1;36(3):849–51.

39. Blagus R, Lusa L. Class prediction for high-dimensional class-imbalanced data. BMC Bioinformatics. 2010 Oct 20;11(1):523.

40. Chen C. Using Random Forest to Learn Imbalanced Data.

41. Solving Class imbalance problem in CNN | Medium | AI Graduate [Internet]. [cited 2023 Dec 13]. Available from: https://medium.com/x8-the-ai-community/solving-class-imbalance-problem-in-cnn-9c7a5231c478

42. Jeong DH, Kim SE, Choi WH, Ahn SH. A Comparative Study on the Influence of Undersampling and Oversampling Techniques for the Classification of Physical Activities Using an Imbalanced Accelerometer Dataset. Healthcare (Basel). 2022 Jul 5;10(7):1255.

43. Fujiwara K, Huang Y, Hori K, Nishioji K, Kobayashi M, Kamaguchi M, et al. Over- and Under-sampling Approach for Extremely Imbalanced and Small Minority Data Problem in Health Record Analysis. Front Public Health. 2020 May 19;8:178.

44. Brownlee J. Undersampling Algorithms for Imbalanced Classification [Internet]. MachineLearningMastery.com. 2020 [cited 2023 Dec 13]. Available from: https://machinelearningmastery.com/undersampling-algorithms-for-imbalanced-classification/

45. Gnip P, Vokorokos L, Drotár P. Selective oversampling approach for strongly imbalanced data. PeerJ Comput Sci. 2021 Jun 18;7:e604.